

**Economic and Social Council**Distr.  
GENERALE/CN.3/1983/2  
2 August 1982

ORIGINAL: ENGLISH

---

STATISTICAL COMMISSION  
Twenty-second session  
7-16 March 1983  
Item 3 (a) of the provisional agenda\*

## SPECIAL ISSUES: ADMINISTRATIVE RECORDS

Use of administrative records for statistical purposesReport of the Secretary-General

As requested by the Statistical Commission at its twenty-first session and its Working Group on International Statistical Programmes and Co-Ordination at its ninth session, the present report provides a review of recent developments and issues in the use of administrative records for statistical purposes. It presents summary information on the experience of Canada, France, Israel, the Netherlands and the United States of America, discusses a number of issues that arise in the use of administrative records for statistical purposes and offers proposals and priorities for further work at the national and international level.

In the review of national experience, the extensive and varied contributions of administrative records, particularly computerized administrative records, are illustrated. Among the problems and approaches discussed are co-ordination and compatibility of coverage, concepts, classifications and reporting units, logical and physical integration of data sources through various kinds of matching and institutional and managerial arrangements for policy and technical co-ordination, quality control and the protection of privacy.

For future work at the national level, overall statistical co-ordination and the development and broad application of information technology are emphasized. At the international level, the need for handbooks and technical reports on the development and operation of specific kinds of administrative records systems is considered from the statistical viewpoint. Points for discussion by the Commission are included (para. 71).

---

\* E/CN.3/1983/1.

PLEASE RETURN TO THE SECRETARIAT OF THE UNITED NATIONS

CONTENTS

	<u>Paragraphs</u>	<u>Page</u>
INTRODUCTION .....	1 - 6	3
I. CURRENT SITUATION IN SELECTED COUNTRIES .....	7 - 20	4
A. Canada .....	10 - 11	5
B. France .....	12 - 13	5
C. Israel .....	14 - 15	6
D. Netherlands .....	16 - 17	7
E. United States of America .....	18 - 19	7
F. Discussion .....	20	8
II. ISSUES RELATED TO THE USE OF ADMINISTRATIVE RECORDS FOR STATISTICAL PURPOSES .....	21 - 48	8
A. Coverage .....	23 - 27	8
B. Concepts, classifications and reporting units .....	28 - 31	9
C. Logical and physical integration .....	32 - 36	10
D. Co-ordination between administrative and statistical systems .....	37 - 39	11
E. Other problems .....	40 - 48	12
III. PROPOSALS AND PRIORITIES FOR FURTHER WORK .....	49 - 70	14
A. Work at the national level .....	50 - 63	14
1. Conceptual and technical issues .....	50 - 53	14
2. Organizational and institutional issues .....	54 - 63	15
B. Future role of the United Nations Statistical Office .....	64 - 70	17
IV. POINTS FOR DISCUSSION .....	71	18
<u>Annex.</u> QUESTIONNAIRE ON THE USE OF ADMINISTRATIVE RECORDS FOR STATISTICAL PURPOSES		

BIBLIOGRAPHY

## INTRODUCTION

1. At its twenty-first session, the Statistical Commission "(a) Decided that at the next several sessions, agenda items should include specific important aspects of future development work" 1/ and at the same session, when considering the draft provisional agenda for its twenty-second session, selected "the use of administrative records for statistical purposes" as a subject for consideration as a special issue". 2/ At its ninth session, the Commission's Working Group on International Statistical Programmes and Co-ordination gave instructions on how the document should be prepared and on its contents (E/CN.3/1983/24, paras. 6-8).

2. The present report has, so far as possible, been prepared in line with those instructions and with a view, also expressed by the Commission at its twenty-first session, to the value of future-oriented discussions. 3/ A short introduction describes (a) the background of the present document; (b) some of the reasons why more attention is being given to administrative records as a source of statistics in a number of countries; and (c) the role of administrative records in balanced and integrated statistical development. The document then presents a limited and illustrative review of the current situation in the use of administrative records for statistical purposes, based on country responses to a United Nations Statistical Office study of national practices (section I). Based on the study of national practices, other relevant work by the Statistical Office and by others, section II contains a discussion of a number of the major issues that have arisen in the use of administrative records as a source of statistics. The document concludes with a discussion of proposals and priorities for further work at the national and international level (section III). A number of points for discussion by the Commission are indicated (section IV).

3. Both the review of current national experience and of issues are intended to stimulate discussion, particularly on matters raised in section III. Although portions of the document may eventually be included in technical reports or other documents of the United Nations, the present document has not been prepared with publication as its objective. Rather it is seen primarily as a starting point for further work in this important, evolving and complex area.

4. The use of administrative reporting systems as a source of statistics goes back to the origins of official statistics. Indeed, the modern population census was preceded in many countries by an administrative census carried out for taxation or other administrative purposes. Moreover, major portions of routine official statistics in most countries have always come primarily or completely from administrative records (for example, external trade statistics, income and tax statistics, vital statistics).

5. Over and above the traditional and continuing use of administrative records for statistical purposes, there is renewed attention in many countries to the importance of administrative records as a statistical source. A number of factors account for this. They include: (a) technological developments in the electronic data processing field, which permit rapid and massive storage, retrieval and manipulation of computerized administrative records, or parts thereof, at increasingly lower cost; (b) the rapidly increasing cost of censuses and sample

surveys in a number of countries, which is often aggravated by the fact that the full cost of these inquiries must be funded out of an identifiable statistical budget; (c) the increased emphasis on regional and local planning as a necessary complement to national planning, with the consequent demand for disaggregated data far beyond the capacity of most sample surveys; (d) the expanding role and function of national and local governments in many countries; and (e) the efforts to minimize the burden placed on individuals and enterprises by overlapping governmental information-gathering efforts. In the language of Nathan and Sicron, both "supply" and "demand" factors have contributed to the renewed interest by statisticians, users and policy makers in the opportunities that administrative records present for major advancements in statistical development. 4/

6. Even in a document, such as the present one, in which the contribution that administrative records can make to statistical development is stressed, it is important to bear in mind that administrative records must be seen as but one part of a comprehensive and integrated statistical system. As is stated in the Handbook of Statistical Organization:

"A balanced programme for the improvement of national statistics involves the use of censuses, sample surveys and administrative records. In the long run these three sources of statistics are complementary to a considerable extent; each has some advantages and suffers from some limitations. Therefore the full development of one source does not render the other two sources superfluous." 5/

A recent United States National Academy of Sciences report on data collection reached a similar conclusion:

"There is no one best source for obtaining fertility and mortality data and no one best design for applying any given data collection method. For this reason, if all users ... are to be effectively served, attention and resources must be directed to a balanced program of upgrading census, registration, and survey capabilities and activities within developing countries. This balanced program will vary among countries because priorities, time constraints on data needs, and planning horizons differ." 6/

#### I. CURRENT SITUATION IN SELECTED COUNTRIES

7. As part of its instructions for the preparation of the present document, the Working Group on International Statistical Programmes and Co-ordination considered that "information on country experience should be obtained through a well-structured inquiry to countries" (E/CN.3/1983/24, para. 8). Drawing on a draft prepared by Data for Development, the Statistical Office prepared the "Questionnaire on the Use of Administrative Records for Statistical Purposes" and circulated it to a diverse group of 31 developed and developing countries in December 1981. The Questionnaire is reproduced in the annex to the present document.

8. As of 24 March 1982, the cut-off date used in preparing the document, substantive responses had been received from five countries (Canada, France, Israel, the Netherlands and the United States of America). In addition, one country indicated it was not in a position to reply, and another indicated that it would not be able to reply in the time provided. No reply of any kind was received from the other 24 countries. 7/

9. Clearly, the five countries which provided substantive responses to the study are among the more statistically developed Member States; nor do they adequately represent the full range of experience and circumstances encountered even among the statistically more developed countries. Thus, it would be unwise to draw many general conclusions from their responses. For this reason, the present section of the document simply contains a summary of the range of administrative records and statistical outputs and uses identified by the five responding countries (that is, essentially a summary of the responses to items 1 and 2 of the questionnaire) to give at least some indication of the diversity and potential of the field. The replies of the five responding countries relating to items 3 to 6 of the questionnaire were used in preparing sections II and III.

#### A. Canada

10. In response to the first question in the study, Canada identified 23 administrative record systems that "are currently computerized and are in varying degrees used for statistical purposes (i.e., some are used more heavily than others)". Of the 23, 4 pertain to vital records, 5 pertain to health records, 3 to the judicial system, 5 to various social security programmes, 3 to education, 1 to agriculture and 2 to the tax system.

11. In response to the second question, Canada identified 52 separate publications listed in the 1981 edition of the Statistics Canada Catalogue as examples of statistical outputs based on administrative records. It was also noted that in addition to the publications, certain data files are available in "machine readable form so that individual users can obtain tabulations according to their own specifications".

#### B. France

12. France identified over 110 record systems used for statistical purposes in 19 subject fields. It characterized the systems according to the extent of participation of the regular national statistical services in their design and management as follows:

(a) Mainly aggregate data used from administrative compilations for statistical purposes;

(b) Administrative forms used as the direct source of data for statistical purposes, in whose design and compilation the statistical services have some degree of responsibility;

/...

(c) Continuous record systems based on unit identifiers designed largely for management purposes but also used for statistics;

(d) Integrated data systems designed to meet a variety of administrative, managerial and statistical requirements.

France listed 33, 60, 30 and 17 record systems respectively of each type. It also noted several cases in which identifier-based systems have been merged to improve the scope and coverage of these sources for statistical purposes, most notably in the field of enterprise statistics, where the national statistical office maintains a national registry of enterprises and establishments. The extent to which each reporting system is computerized was not indicated in the reply.

13. Two types of restrictions on the the use of administrative records for statistical purposes were noted by France. First, according to the principle of final use adopted in France, administrative forms cannot be used to collect information other than that required for specific legislative purposes. Secondly, among the various branches of the government, there are various legal requirements for maintaining confidentiality. In each case where the statistical office wishes to compile statistics from the records of other services, special arrangements must be made which, if they are possible at all, are subject to various restrictions. These are all the more difficult when it is desirable to use two sources together.

#### C. Israel

14. Israel indicated that it uses administrative records extensively for statistical purposes. Major portions of the statistical system covering population, vital statistics, migration, international trade, labour and wages, transportation, the judiciary, education, health and welfare are based completely or primarily on the use of administrative records. Some examples of the major administrative record systems described in the response are (a) the Israeli Population Register; (b) the National Insurance Institute Employer's File; and (c) the files of discharges from hospitals.

15. Several factors, as noted in the response of Israel, contribute to the widespread use of administrative records for statistical purposes. First, a personal identity number is commonly used, as a unique identifier, in a variety of administrative record files, thus enhancing the potential for record linkage and data integration. Secondly, the introduction of computers throughout the governmental administration as well as in private establishments has substantially contributed to the quality, ease and timeliness in processing massive administrative files. Thirdly, the application of standard practices and classifications has enabled the direct processing of records reported by different agencies or localities.

D. Netherlands

16. Numerous statistical uses of administrative records were reported by the Netherlands. Its response was organized in terms of the most important administrative reporting systems operative in the following subject-matter areas: population, health, building, industry, education, the judiciary, income and wealth, external trade, social security and wages, public expenditure and finance, agriculture and transport. Statistics for several of the fields are based almost entirely on administrative records, for example, population, income and wealth, external trade, public expenditure, the judiciary, building and industry. Several other fields, for example, social security, labour and wages, health and transport, make extensive use of administrative records, in conjunction with records of semi-governmental agencies (labour offices, trade association etc.) for deriving important statistics of unemployment, paid working days, total pay-roll, work accidents and disability, traffic accidents, transport of persons or goods by rail and inland navigation of vessels.

17. The extent of computerization is reported to vary from one field to another. Many municipalities have computerized population registers, and the system yields numerous monthly series and some annual data. In the Netherlands, a unique personal number exists only for inhabitants of municipalities with computerized population registers (about 8 million persons). In most other fields, the records do not contain a unique identifier. This situation is seen as making the task of record linkage more difficult and limiting the joint use of the administrative files.

E. United States of America

18. In a recent study of statistical uses of administrative records, carried out under the auspices of the then Office of Federal Statistical Policy and Standards, more than 20 administrative record files were examined in 10 federal agencies containing information on individuals and 10 administrative record files in 6 federal agencies that refer mainly to businesses. The United States response was based on the results of the study.

19. Most administrative record files for individuals are reported to contain relatively little information on population characteristics, compared to the decennial population census. Further, the files often cover only limited segments of the population. Most files, however, contain the social security number of individuals, opening the possibility for linking files over time or merging information from different files in order to increase the coverage of individuals or the number of characteristics identified for particular individuals. However, the absence of social security numbers in the population census records makes it difficult to integrate information from the census with information from administrative records. The coverage of businesses in administrative record systems is reported to be more complete than for individuals, and administrative files on businesses, in addition to generating numerous statistical series directly, provide the framework for conducting many censuses and surveys. On the other hand, many administrative record files pertaining to businesses do not contain the information required for deriving reliable detailed geographical and industrial data.

## F. Discussion

20. The summarized responses of Canada, France, Israel, the Netherlands and the United States given in paragraphs 10-19, above, vividly illustrate the extensive contribution that administrative records in general, and computerized administrative records in particular, are already making to statistical outputs and capacities in those countries. What is equally apparent from the summaries, and even more so from the detailed material generously provided, is the diversity and complexity of the field. Even within this comparatively homogeneous group of countries, there is great variation in what they considered to be an administrative reporting system, how they organized their responses and how they defined statistical outputs and uses. If a useful study of national experience is to be made, considerably more time will be required than was available for the preparation of the present document. Time is needed not only to give countries an adequate interval in which to respond, but also for an evaluation and distillation of the country responses and, where required, for correspondence with countries about certain aspects of their responses.

## II. ISSUES RELATED TO THE USE OF ADMINISTRATIVE RECORDS FOR STATISTICAL PURPOSES

21. There can be no question about the desirability of making maximum use of administrative records for statistical purposes. A multitude of administrative record systems exist in each country, touching upon most facets of economic and social activity and collecting vast amounts of information concerning enterprises and individuals. When such systems function well, both in an administrative and a statistical sense, they can economically generate timely data that are free from sampling error, even when highly disaggregated. Furthermore, the institutional continuity of some administrative record systems is conducive to the generation of comparable time series that may cover decades or even centuries. Finally, there are operational and policy considerations that make administrative record systems a highly attractive statistical source (see para. 5 above).

22. The purpose of the present section is to review some of the obstacles to the expanded use of administrative records for statistical purposes. It should be stressed, however, that though focusing on the problems encountered in using administrative records, there is no question of their importance as a statistical source. A review of the problems is a necessary step in the further development of administrative record systems as a source of statistics. The specific issues discussed were, for the most part, suggested by Data for Development.

### A. Coverage

23. The term "coverage problems" is used to refer to two quite distinct difficulties. The first problem refers to the fact that any specific administrative record system covers only those units (individuals, transactions, entities etc.) provided for under the laws, regulations or rules governing the



system. Statistics directly produced from these records will refer to a well-defined, but restricted, population (for example, all businesses employing five or more persons or all live births surviving 72 hours). The restricted populations involved are usually specific to each administrative system and, in the case of even moderately decentralized systems, the coverage rules may vary within the country.

24. Such coverage restrictions, although integrally related to the administration of the specific law or regulation in question, usually serve as a barrier to the use of statistics derived from any given administrative system for more general purposes. To serve these more general purposes, it is often essential that data from two or more administrative systems be merged or that data from one or more administrative systems be used jointly with statistics from a census or one or more sample surveys. Unfortunately, the highly specific nature of the coverage restrictions involved often makes this very difficult.

25. The second type of coverage problem arises when some units that should have been covered by a specific administrative reporting system are omitted, for example, the omission of enterprises or vital events from official registers because no one reported them to the appropriate officials. This type of coverage problem occurs, of course, in all countries, but it is particularly common in developing countries. In many developing countries, the problem is encountered in a broad range of administrative reporting systems and is also characterized by relatively high omission rates.

26. Both statistical and administrative personnel can easily agree about the existence of the first type of coverage problem, although they often find it considerably harder to agree on whether anything can or should be done about it. With respect to the second type of coverage problem, the situation is somewhat different. Some administrative personnel may find it difficult, at least initially, to accept the existence of substantial omissions and it is often the statistician who first brings to their attention evidence of the size of the problem. However, once the fact of the problem is agreed upon, it is reasonable to assume that administrative and statistical personnel have a common interest in reducing the omissions.

27. Finally, although the two types of coverage problems are quite distinct, there can be, at least in certain circumstances, a trade-off between them. For example, by restricting an administrative reporting system to a more limited population, one may be able to reduce the second type of coverage problem. Whether or not any real gain is achieved thereby depends on the circumstances.

#### B. Concepts, classifications and reporting units

28. Among the major obstacles to the better use of administrative records for statistical purposes are the concepts, definitions and classifications used in administrative systems. Unless steps are taken to overcome these problems, the concepts, definitions and classifications usually lack the generality needed to serve many important statistical purposes.

29. Just as the population covered by most administrative record systems is defined in the highly restricted and specific language of the law or regulation governing the system, many of the concepts, definitions and classifications have similar origins. However, many other concepts, definitions and classifications used in administrative reporting systems have less substantial origins that are often based on outdated theory, an overly narrow definition or perception of the phenomenon, administrative convenience or ignorance etc.

30. Generally, concepts, classifications etc. that are incidental to the laws or regulations governing an administrative reporting system are easier to alter than those that are specified in the laws or regulations or are central to the operation of an administrative system. But regardless of their source, the obstacles posed can be extensive and challenging to overcome. Indeed, a good portion of ongoing national and international work in statistical methodology is aimed at overcoming or coping with the fact that the concepts, definitions, classifications and reporting units used in the generation of many data sets are not well suited to the uses to which the data are put.

31. Two examples may be cited. Substantial portions of Studies in the Integration of Social Statistics: Technical Report contain a discussion of the problem of harmonizing concepts, definitions, classifications and reporting units from a statistical viewpoint. 8/ The reports of two recent United Nations meetings, one on external trade 9/ and the other on national accounts 10/, attest to the complex issues that must be addressed and compromises reached if administrative reporting systems are to serve statistical needs more adequately in the future than they have done in the past.

### C. Logical and physical integration

32. One of the major advantages of using administrative record systems for statistical purposes is the benefit derived from merging information obtained at different times or from different administrative systems for the same unit (for example, person, institution, transaction, event).

33. The merging of the information may be for administrative purposes, for statistical purposes or for both. If the merging is to serve an administrative end (for example, merging information on all arrests for the same individual), an unambiguous linkage between two or more records for the same units must be established. Such a linkage requires an "exact matching" procedure. 11/ Furthermore, the exact matching procedure must be designed to yield no erroneous matches, if administrative confusion and grave injustice, at least in the case of sanctions, are to be avoided. Exact matching, with no erroneous matches, implies for most countries and most administrative record systems unique identifiers, for example, numbers for persons or codes for business establishments. 12/ In theory, if a linkage can be established for administrative purposes, the merged file should be available for statistical purposes, although problems of timing and access frequently arise (see paras. 43-44 below).

34. An exact matching procedure of this type, that is, one with no (or few) erroneous matches (and some reasonable bounds on the level of erroneous

/...

non-matches) is often expensive and time-consuming. Furthermore, a system of identifiers less ambiguous than names and addresses must be accepted and known by the concerned public (individuals, establishments etc.), while knowledge and use of such special identifiers places an added burden on form fillers. Even more of a problem in a number of countries is the political acceptability of special identifiers for persons. This problem is part of the larger issue of the political acceptability of record linkage for administrative purposes. Again, the issue is complex: all countries will consider some kinds of linkage acceptable, but few will consider all kinds acceptable.

35. Over and above the linkages between administrative record systems that are carried out for administrative purposes, there are those that are carried out on administrative records for statistical purposes. As described in The Development of Integrated Data Bases ..., there are two procedures for record linkage for statistical matching: "exact matching" and "statistical matching". <sup>13/</sup> Exact matching for statistical purposes - in contrast to matching for administrative purposes - need not place so high a priority on achieving no erroneous matches. Here, more latitude is possible in defining an optimum matching rule in terms of the expected level of both erroneous matches and erroneous non-matches or even other aspects of the linkage process. <sup>14/</sup> As a result, the need for a unique, special identifier is less critical, the cost of matching may be lower and the special burden placed on the public is reduced. Nevertheless, at least in some countries, neither the law nor public opinion makes much of a distinction between exact matching for statistical purposes and that for administrative purposes.

36. Statistical matching as a procedure for merging information from two or more data sources for statistical purposes has already been extensively treated. <sup>15/</sup> Suffice it to say that, although statistical matching overcomes most of the public policy obstacles to exact matching (as do other essentially analytical approaches to information merger involving regression techniques or other forms of modelling), all these approaches are highly dependent on the resolution of the coverage and conceptual problems discussed in paragraphs 23-31 above.

#### D. Co-ordination between administrative and statistical systems

37. A series of factors revolving around the issue of co-ordination between the administrative and statistical systems is perhaps the central obstacle to the greater use of administrative records for statistical purposes. Administrative systems, although they serve a common end of good governance, are highly specific and varied in their coverage, instrumentalities and operations. They also may function at the national, provincial or local levels. Finally, while some changes in the organization or operation of an administrative record system can be achieved by the decision of a single, relatively junior civil servant, others may require action by the national legislature or even a change in the country's constitution. These factors contribute to the problems summarized in paragraphs 23-31 above and to the resistance of these problems to change.

38. Naturally, when a large number of administrative and statistical agencies are involved, commitment to change may be hard to achieve or, if achieved, it may be

difficult to translate the commitment into actual progress. The success of efforts directed at improved co-ordination between administrative and statistical systems depends upon (a) executive-level support from the administrative agencies and statistical services involved; (b) knowledgeable technical inputs from the administrative, statistical and data processing sides; and (c) resources for staff to develop new plans and procedures and for experimentation and pilot testing. Furthermore, it must be recognized that while a surprising number of relatively simple changes in administrative systems can yield many benefits in terms of better data and better administration, most improvements take long to achieve and often involve complex trade-offs. Thus, the effort must be seen as a long-term one, involving continuing commitment and continuing activities.

39. The need for continuing attention to the issue of co-ordination between administrative and statistical systems also arises, because the administrative activities of governments are subject to constant change. Clearly, the changes may have serious unintended consequences for the statistical system. Once again, the diversity among administrative systems compounds the problem. Finally, while a co-ordinating mechanism may influence changes arising from within the administrative system, it is very rare to have statistical consequences taken into account when making legislatively or constitutionally based changes in administrative systems.

#### E. Other problems

40. Several other problems arise in the use of administrative data for statistical purposes. A few of the problems are reviewed in the following paragraphs.

41. Statistics derived from administrative records, like statistics derived from direct inquiries, are subject to errors of reporting, recording, processing and compilation. Two types of errors may usefully be distinguished: errors pertaining to items of information directly relevant to the administrative activity or process in which the record is used and errors that pertain to other items of information.

42. In the case of items of information essential to the administrative process, the over-all level of error is likely to be low - assuming, of course, the administrative system involved is well run - with the bias component of total error probably dominant. With regard to other items of information, error rates may be very high, even in a well-run administrative system. This is particularly so when quality control efforts either ignore the statistical, as opposed to the administrative, uses of the records or have insufficient resources.

43. The timing and the nature of access that statistical personnel have to administrative record systems may also affect the usefulness of the record systems as a source of current statistics. Sometimes the administrative processes involved, or the way in which the record system operates, preclude early access to the administrative file for statistical work. Although there are several ways to overcome the obstacles (for example, an advance tabulation sample for statistical purposes, the creation of a separate statistical record at an early stage of the administrative process or even a redesign of certain features of the operating

procedures used in the administrative system), such approaches usually require additional resources in order to be planned and implemented successfully. The situation is even more complex when the record system is subject to updating, whether periodic or intermittent. If the file is large, updating it can be an immense task subject to unpredictable delays.

44. Certain administrative activities are considered confidential and statistical personnel may be permitted access to the record system only at certain stages in the process or only after certain information is suppressed by the concerned administrative agency. Such barriers to access may adversely affect the statistical usefulness of an administrative record system because, in these circumstances, it is difficult to introduce meaningful quality control and broader evaluative procedures. Restrictions on the access of statistical personnel to the unedited information in administrative files may also create problems, because the editing requirements of the administrative and statistical organizations frequently differ. One goal of improved co-ordination between administrative agencies and the national statistical system (see paras. 37-39 above) is to reduce the barriers, or at least their impact, while taking into account the legitimate concerns of the administrative agencies involved, including confidentiality requirements.

45. Finally, it must be noted that administrative record systems pertaining to individuals are usually quite limited in the extent to which they provide information on important socio-economic classification variables. The variables are often seen as irrelevant to the administrative task at hand and thus their collection is viewed as imposing an unjustified burden on the public and on administrative personnel. Furthermore, questions can properly be raised about the level of response error (both bias and response variance) expected in information on socio-economic topics obtained in many administrative settings. Nevertheless, the classification variables (for example, educational attainment, income, occupation) are essential for many policy and intersectoral uses of statistics.

46. To overcome the problem, it is particularly important to be able to use information from an administrative record system jointly with that obtained from other sources (that is, censuses, sample surveys and other administrative record systems). This, in turn, places special emphasis on the need for work on harmonizing concepts, definitions, classifications and reporting units, as described in paragraphs 28-31 above. In particular, where unique identifiers are not available or exact matching is not feasible or desirable for other reasons (see paras. 32-36 above), the use of common age-group classifications and geographical identifiers is of central importance to the use of statistical matching or regression techniques.

47. With respect to age-group classifications, information on a recent review of national practices was contained in the document before the Statistical Commission at its twenty-first session, "Draft guide-lines on age-group classifications" (E/CN.3/550). International guidance on the subject is given in Provisional Guidelines on Standard International Age Classifications 16/. Both underscore the fact that, while there is considerable scope for improving the current situation, a host of different age classifications are currently in use in most countries. Some of the variations are dictated by legislative requirements or other special circumstances, but many have no apparent rationale.

/...

48. The situation is even more difficult with respect to the standardization of geographical identifiers or the use of common geographical classifications at the national level. <sup>17/</sup> Geographical identifiers and in many administrative reporting systems, at least the levels of greater disaggregation are often specific to a given administrative system and are frequently embodied in the governing legislation or regulations. Unless a country has a comprehensive and flexible computerized geocoding system, the provision of common identifiers is apt to be an error-prone and costly process.

### III. PROPOSALS AND PRIORITIES FOR FURTHER WORK

49. The present section offers, for discussion purposes, a number of suggestions for further work. Sub-section A is concerned with proposals and priorities for work at the national level, while subsection B is concerned with work at the international level, focusing on the work of the United Nations Statistical Office.

#### A. Work at the national level

##### 1. Conceptual and technical issues

50. As indicated above, particularly in paragraphs 23-31 and 46-48, a number of conceptual, definitional and related problems need to be addressed if administrative record systems are to serve more adequately a broad range of statistical uses. Various aspects of national work on these issues have been described in Handbook of Statistical Organization <sup>18/</sup> and Studies in the Integration of Social Statistics: Technical Report. <sup>19/</sup> More specialized information and guidance may be found in Methodology and Evaluation of Population Registers and Similar Systems <sup>20/</sup>; International Trade Statistics: Concepts and Definitions <sup>21/</sup>; Principles and Recommendations for a Vital Statistics System <sup>22/</sup>; "The harmonization of statistical classifications: report of a meeting of an Expert Group" (ST/ESA/STAT/78); Guidelines for the Collection of Statistics on Literacy Programmes <sup>23/</sup>; Balance of Payments Manual <sup>24/</sup>. In addition, the following documents submitted to the Statistical Commission at its present session illustrate the scope and complexity of ongoing work on concepts, definitions, classifications and reporting units having substantial relevance for statistics based on administrative reporting systems: E/CN.3/1983/5, E/CN.3/1983/6, E/CN.3/1983/7, E/CN.3/1983/8, E/CN.3/1983/9, E/CN.3/1983/10, E/CN.3/1983/12, E/CN.3/1983/15, E/CN.3/1983/17.

51. National efforts designed to resolve conceptual, definitional and related problems encountered in the use of administrative records for statistical purposes should be carried out as part of the over-all effort of national statistical co-ordination; indeed, in a number of countries, the efforts will be the most important, complex and resource-intensive aspects of such co-ordination and integration activities. National priorities for this work can only be established in the light of the needs, problems, opportunities and circumstances of individual countries.

52. With respect to technical issues, the likely future direction of most aspects of work at the national level is reviewed in the document entitled "Implications of new developments in information technology", which has been submitted to the Commission at its present session (E/CN.3/1983/3). The general trends in information processing outlined therein are of particular relevance to the compilation and use of statistics derived from administrative records. Therefore, only one supplementary point will be made here.

53. It must be recognized that the increased availability of low-cost computing power through mini-computers and micro-computers may, in the short run and particularly for developing countries, make it more difficult to extract consistent statistics from administrative record systems. As national data processing activities become increasingly decentralized by agency and unit within agency and by administrative level, there will be a natural tendency for greater decentralization and variation in the use of standard identifiers, formats, file structures and classifications as well as in the substantive content of the registers themselves. Any long-term national effort aimed at making greater use of administrative record systems as a source of statistics will need to identify and provide specific technical and institutional means to offset this tendency.

## 2. Organizational and institutional issues

54. The greatest obstacles to using administrative records more effectively and efficiently for statistical purposes do not appear to be conceptual or technical in nature; rather they are organizational and institutional. Therefore, these issues cannot be avoided if countries wish to make better use of administrative record systems as sources of statistics.

55. Among the organizational and institutional issues that need to be resolved are the respective roles and responsibilities of the administrative, statistical and data processing personnel and agencies involved, methods of achieving co-ordination, the provision of needed resources and the timely and adequate access of statistical personnel to administrative files.

56. Two of the issues were discussed above, namely co-ordination (paras. 37-39) and access (paras. 43-44). Resources, although already mentioned several times, is an issue which deserves special attention.

57. Simply stated, administrative record systems are not a free source of data. Of course, the administrative reporting system is functioning at no cost to the statistical budget, but when one wants to use such a reporting system for statistical purposes, added resources are required. The extent to which the cost of the added resources is borne by the administrative or statistical services or shared between them varies among reporting systems, levels of government and countries. What is universally true is that if insufficient resources are provided the quality or timeliness, and often both, will suffer. Thus, it is essential to take additional resource requirements into account when considering the comparative costs of alternative data sources or planning to use a particular administrative reporting system as a statistical source.

/...

58. The added resource needs include resources for, inter alia, statistical planning and co-ordination, conceptual work, special statistical forms, quality control procedures, training, editing, reformatting and file manipulation, record linkage and adjustments to make the data useful. Despite the added resources involved, an administrative record system may be less costly than an alternative data source. However, one cannot be certain until added resource needs have been considered.

59. In considering the relative roles and responsibilities of administrative, statistical and data processing personnel and agencies, it is important to note that in recent years, in large part because of a number of the developments noted in the document on information technology (see para. 52 above), the system of interrelationships on which much of the statistical work had rested is undergoing a process of fundamental re-adjustment in many countries. Drawing attention to this process, Malinvaud has written "... there is a paradoxical aspect since statistics, which has long set the pace in data processing, is now almost bringing up the rear on certain points. For some time now statisticians have not been the only ones to manipulate data in great numbers. Perhaps they, however, want boldness in making use of the technical developments achieved in other sectors." 25/

60. Malinvaud noted that one major consequence of these developments was the need for "a re-examination of the co-ordinative role played by national offices of statistics. The point is not to change radically this role, but merely to re-orientate it gradually ..." 26/ He gave a number of telling examples of the importance and complexity of this effort, all arising from work on the use of administrative records for statistical purposes.

61. The issues which Malinvaud addressed also manifest themselves in many other countries. Frequently, statisticians and national statistical offices have remained uninvolved when others proceeded to develop "national information systems", "national information policies" or "management information systems". Such systems and policies are attractive to administrators and policy-makers and often receive considerable support. The scope of the systems varies - bibliographical materials, statistics, non-quantitative information - but at the centre is information, largely statistics, derived from administrative record systems. Indeed, the systems and policies may become the principal means by which statistics are obtained from administrative records. Unfortunately, because statisticians were not usually involved in the planning and implementation of the information systems, little or no attention was given to such basics as the use of standard statistical concepts, definitions and classifications, quality-control procedures etc. and the resulting data were of limited use. It is apparent that any effort at further exploiting administrative record systems for statistical purposes should involve national statistical offices in national information policies and systems.

62. Since statistical information constitutes a major (if not the major) part of the national information system, it is logical that the national statistical organization should take a leading role in the development of a national information system and related information policies. Clearly, other governmental agencies such as the planning organization and the national data processing centre,



if one exists, would also be active in the effort but these organizations have already shown a keen interest in these developments in many countries. Generally, national statistical offices have been less active in this area.

63. One approach to the development of a sound national information system and related policies is the establishment of a high-level interagency committee in order to provide over-all policy direction and to ensure the continued co-operation of the various governmental agencies in providing access to their administrative record systems. The high-level policy committee would need to be assisted by one or more technical committees composed of administrative personnel, statisticians, management specialists, computer technicians etc. to deal with the various technical and substantive issues involved. One task of the technical committee would be to ensure that work on concepts, classifications, coding systems, quality-control procedures etc. takes into account the special requirements of the statistical component of such an information system.

#### B. Future role of the United Nations Statistical Office

64. The methodological work of the international agencies, in general, and the United Nations Statistical Office, in particular, in support of the use of administrative records for statistical purposes will need to proceed along two parallel lines. The first involves general approaches to the topic (such as the present document) or generally applicable methodologies (such as documents dealing with various aspects of data processing or record linkage). The second line involves work of relevance to specific types of administrative reporting systems.

65. There, of course, remains considerable scope for the first line of work to continue. The study of national practices (see paras. 7-20 above) could be amplified and completed. A general technical report could be prepared, based on the present document and on the discussion by the Statistical Commission and similar discussions by a number of regional bodies of statisticians. Several technical reports could be prepared on various aspects of data processing that are of particular relevance to the use of administrative records for statistical purposes. Technical reports on other general issues such as record linkage, confidentiality policies, sampling or quality-control procedures of special relevance to administrative records could be prepared. All of the outputs mentioned in the present paragraph would appear to be useful and, perhaps, the first two or three mentioned are essential at this stage. However, it is clear that there are limits to what can be achieved through general reports and general methodologies.

66. Of equal importance is methodological work of relevance to specific kinds of administrative record systems. Such work, whether in the form of handbooks, technical reports, draft recommendations or guidelines, draft international classifications etc., has been a major focus of the methodological work of the Statistical Office and is expected to remain so (see para. 50 above, as well as E/CN.3/1983/29).

67. In considering the relevance of more specialized methodological work to the improved use of administrative records for statistical purposes, two points should be kept in mind. First, although much of the work is substantive in nature, most of the outputs have multisectoral uses. For example, the International Standard Industrial Classification of all Economic Activities was developed in the context of economic statistics, but it is widely used, at least at the first-digit level, in social and demographic statistics. Secondly, much of the work is equally relevant for generating better data both from one or more administrative record systems and from censuses and surveys.

68. However, if better use is to be made of administrative record systems, it would seem essential also to begin the preparation of a series of handbooks and technical reports focused on the development and operation of specific kinds of administrative record systems from a statistical viewpoint. A few obvious examples may be cited: customs systems, tax registers, population registers; the list can easily be extended.

69. That much useful and even essential work remains to be done is apparent. What is less clear is whether, to what extent or how priorities should be adjusted to accommodate some or all of the work. With a view to a balanced approach to statistical development, that is, one involving censuses and sample surveys as well as administrative record systems (see para. 6 above), one would call for a selective and marginal adjustment of priorities to accommodate, over the next two to four years, some of the outputs described in paragraphs 65 and 68, while continuing most of the methodological work mentioned in paragraphs 66-67.

70. In view of the wide interest in the subject expressed by a number of individuals, Member States, regional bodies of statisticians and non-governmental organizations, such as Data for Development, it may be possible to plan for a co-ordinated development of related documentation so that key gaps could be filled, at least provisionally, as rapidly as possible.

#### IV. POINTS FOR DISCUSSION

71. The Commission may wish to:

(a) Comment on the advantages and limitations of using administrative records for statistical purposes in the light of the national experience and the substantive, technical and policy issues reviewed above;

(b) Consider the proposals and priorities for further work at the national and international level (see paras. 49-70).

Notes

1/ Official Records of the Economic and Social Council, 1981, Supplement No. 2 (E/1981/12), para. 261 (a).

2/ Ibid., para. 267.

3/ Ibid., para. 250.

4/ Gad Nathan and Moshe Sicron, "The use of registers and administrative files for preparing and improving statistical data". Invited paper to the forty-third session of the International Statistical Institute, Buenos Aires, 30 November-11 December 1981.

5/ United Nations publication, Sales No. E.79.XVII.17, chap. IX, sect. A.

6/ Committee on Population and Demography, Panel on Data Collection, "Collecting Data for the estimation of Fertility and Mortality", report No. 6 (Washington, D.C., National Academy Press, 1981), p. 42.

7/ Subsequently, substantive replies were received from eight countries (Australia, Brazil, Finland, Hungary, Japan, Mexico, Sweden and the Union of Soviet Socialist Republics) and an additional reply was received from Canada. Furthermore, a very useful set of papers prepared by several members of the Conference of European Statisticians became available in connexion with a discussion of this topic at the thirtieth plenary session of the Conference, in June 1982. (The papers carry the symbols CES/461 and Add.1-Add.6.)

8/ United Nations publication, Sales No. E.79.XVII.4, particularly Part two, paras. 11-40 and Part three, paras. 1-26.

9/ "Report of the Expert Group Meeting on the Third Revision of the Standard International Trade Classification" (ESA/STAT/AC.14/3).

10/ "Report of the Expert Group Meeting on the Review and Development of the United Nations System of National Accounts (SNA)" (ESA/STAT/AC.15/8).

11/ See The Development of Integrated Data Bases for Social, Economic and Demographic Statistics (United Nations publication, Sales No. E.79.XVII.14), paras. 83-87.

12/ "Linking population-based and institution-based data for individuals: procedures and experiences", report prepared by the National Central Bureau of Statistics of Sweden for the Third Joint ECE/WHO Meeting on Health Statistics (CES/AC.36/21 - EURO/ICP/HST/010/21), 26 November 1981.

13/ United Nations publication, Sales No. E.79.XVII.14, paras. 83-98.

14/ It may be noted that, in addition to merging information in different files, exact matching is sometimes used to estimate the omission rate - the second type of coverage error described in sub-section A above - in an administrative record system.

15/ The Development of Integrated Data Bases ..., and "Role of macro-data and micro-data structures in the integration of demographic, social and economic statistics: report of the Secretary-General" (E/CN/3/552).

16/ United Nations publication, Sales No. E.82.XVII.5.

17/ "National practices in classifications of size and type of locality and urban/rural areas: report of the Secretary-General" (E/CN.3/551 and Corr. 1).

18/ United Nations publication, Sales No. E.79.XVII.17.

19/ United Nations publication, Sales No. E.79.XVII.4.

20/ United Nations publication, Sales No. E.69.XVII.15.

21/ United Nations publication, Sales No. E.70.XVII.16 (revision in press).

22/ United Nations publication, Sales No. E.73.XVII.9.

23/ United Nations Educational, Scientific and Cultural Organization (Paris 1979) (CSR-E-34).

24/ International Monetary Fund, 4th ed. (Washington, 1977).

25/ E. Malinvaud, "From statistics to data management: the French difficulties", paper presented to the Seminar on Statistical Data Collection and Processing Systems under New Conditions, Moscow, 21-25 September 1981 (CES/SEM.14/R.2), sect. 1.2.1.

26/ Ibid., sect. 2.

Annex

QUESTIONNAIRE ON THE USE OF ADMINISTRATIVE  
RECORDS FOR STATISTICAL PURPOSES

1. Describe briefly the main administrative reporting systems that are now used for statistical purposes, identifying which have been computerized and which have not. Among those which might be included are the following:

- civil registration
- social security
- frontier controls (immigration and tourism)
- health services and conditions
- educational services and qualifications
- business taxes and licences (by branch of economic activity)
- customs documents
- international currency transactions
- public expenditure and finance
- tax data
- land and agricultural surveys and registration
- construction permits
- building registration

2. What achievements have been gained in your country by the use of administrative records for statistical purposes? Describe briefly some of the principal statistical outputs and their uses in the various fields.

3. Identify and describe briefly any administrative routines in which unique identifiers are currently used, for example, a unique identification number for business establishments. Are any of these numbers common to more than one administrative file? If so, do you perform or could you perform record matching on the basis of the identifier and consequently the cross-tabulation of data produced by various administrative routines covering the same basic entities?

/...

4. Describe what problems have been or could be encountered in using your country's administrative reporting systems for statistical purposes. Among these might be noted the following:

- legislative restrictions (including confidentiality)
- administrative regulations
- incompatibility of concepts, classifications, definitions and time reference and recording periods among different administrative records and between administrative records and statistical requirements
- lack of arrangements or authority for administrative co-ordination
- lack of technical resources

5. What improvements would you suggest could be carried out in your use of administrative records for statistical purposes?

6. What activities can the United Nations Statistical Office carry out that might further assist you in the use of administrative records for statistical purposes? (For example, indicate kinds of technical reports that might be helpful.)

BIBLIOGRAPHY\*

- Aziz, Faye, Harriet Orcutt and Linda Del Bene. Social security data files as a resource for health research. In Statistics of income and related administrative record research. Wendy Alvey and Beth Kilss, eds. Washington, DC, Department of the Treasury, Internal Revenue Service, Statistics Division, 1981. pp. 125-128.
- Bachi, R., R. Baron and G. Nathan. Theory of record linkage and applications in Israel. Bulletin of the International Statistical Institute, 42:766-796.
- Bachi, R., and R. Baron. Confidentiality problems related to data banks. Bulletin of the International Statistical Institute, 43:225-241.
- Data for Development. Data and development. Springfield, Va., United States National Technical Information Service, 1978. pp. 274-279.
- Data for Development. Information systems in government: I. Data gathering and organization at central level. Springfield, Va., United States National Technical Information Service, 1976. pp. 261-301.
- Dörnyei, Jozsef. The architecture and organization of the Statistical Data System. In Information systems in public administration and their role in economic and social development. David Eade and John Hodgson, eds. Amsterdam, North-Holland Publishing Company, 1981. pp. 97-111.
- Dumais, Jim and Ray Shadid. Individual statistics of income: advancing the closeout date. In Statistics of income and related administrative record research. Wendy Alvey and Beth Kilss, eds. Washington, DC, Department of the Treasury, Internal Revenue Service, Statistics Division, 1981. pp. 11-20.
- Durkin, Thomas M. and Otto Schwartz. The SOI [Statistics of Income] quality control program. In Statistics of Income and related administrative record research. Wendy Alvey and Beth Kilss, eds. Washington, DC, Department of the Treasury, Internal Revenue Service, Statistics Division, 1981. pp. 33-39.
- France. Institut national de la statistique et des études économiques. Présentation du système statistique français. Paris, 1981.
- Gottlieb, David B. A data base approach for maintaining a registry of establishments. In Information systems in public administration and their role in economic and social development. David Eade and John Hodgson, eds. Amsterdam, North-Holland Publishing Company, 1981. pp. 231-239.

---

\* The references included are supplementary to those cited in the text.

Leclerc, G. Tax administration and economic statistics: some conceptual and policy issues. Invited paper to the forty-third session of the International Statistical Institute, Buenos Aires, 30 November-11 December 1981.

Lieskovsky, Peter. User's aspects of the Integrated Statistical Information System development. In Information systems in public administration and their role in economic and social development. David Eade and John Hodgson, eds. Amsterdam, North-Holland Publishing Company, 1981. pp. 113-123.

Lunde, A. S., et al. The persons number systems of Sweden, Norway, Denmark and Israel. Vital and Health Statistics Series 2, No. 84. Washington, DC, National Center for Health Statistics, 1980.

Norris, Douglas, John Leyes and Nelson Kopustas. The development of social data from administrative records: the Canadian experience. In Statistics of income and related administrative record research. Wendy Alvey and Beth Kilss, eds. Washington, DC, Department of the Treasury, Internal Revenue Service, Statistics Division, 1981. pp. 91-95.

Panel on Data Collection, Committee on Population and Demography. Data collection: a statement for administrators. Report No. 7. Washington, DC, National Academy Press, 1981.

Rapaport, Edmund. Are statisticians entitled to use administrative records? Invited paper to the forty-third session of the International Statistical Institute, Buenos Aires, 30 November-11 December 1981.

Smith, Martha E. and John Silins. Generalized iterative record linkage system. In Statistics of income and related administrative record research. Wendy Alvey and Beth Kilss, eds. Washington, DC, Department of the Treasury, Internal Revenue Service, Statistics Division, 1981. pp. 147-158.

United States. Department of Commerce. Office of Federal Statistical Policy and Standards. Report on statistical uses of administrative records. Statistical Policy Working Paper 6. Washington, DC, 1980.

-----